

# 8



OIPE

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/940,235

DATE: 04/23/2002

TIME: 15:34:20

Input Set : A:\07064-009002.TXT

Output Set: N:\CRF3\04232002\I940235.raw

```

4 <110> APPLICANT: Kumar, Rajesh
5     Sahni, Girish
6     Roy, Chait
7     Rajagopal, Kammara
8     Nihalani, Deepak
9     Sundaram, Vasudha
10    Yadav, Mahavir
13 <120> TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
14     PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
15     CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
16     PROTEIN
18 <130> FILE REFERENCE: 07064-009002
20 <140> CURRENT APPLICATION NUMBER: 09/940,235
C--> 21 <141> CURRENT FILING DATE: 2002-04-09
23 <150> PRIOR APPLICATION NUMBER: 09/471,349
24 <151> PRIOR FILING DATE: 1999-12-23
26 <150> PRIOR APPLICATION NUMBER: IN 3825/DEL/98
27 <151> PRIOR FILING DATE: 1998-12-24
29 <160> NUMBER OF SEQ ID NOS: 28
31 <170> SOFTWARE: FastSEQ for Windows Version 4.0
33 <210> SEQ ID NO: 1
34 <211> LENGTH: 1245
35 <212> TYPE: DNA
36 <213> ORGANISM: Streptococcus equisimilis
38 <220> FEATURE:
39 <221> NAME/KEY: CDS
40 <222> LOCATION: (1)...(1242)
42 <400> SEQUENCE: 1
43 att gct gga cct gag tgg ctg cta gac cgt cca tct gtc aac aac agc      48
44 ile ala gly pro glu trp leu leu asp arg pro ser val asn asn ser
45 1          5          10          15
47 caa tta gtt gtt agc gtt gct ggt act gtt gag ggg acg aat caa gac      96
48 gln leu val val ser val ala gly thr val glu gly thr asn gln asp
49          20          25          30
51 att agt ctt aaa ttt ttt gaa atc gat cta aca tca cga cct gct cat      144
52 ile ser leu lys phe phe glu ile asp leu thr ser arg pro ala his
53          35          40          45
55 gga gga aag aca gag caa ggc tta agt cca aaa tca aaa cca ttt gct      192
56 gly gly lys thr glu gln gly leu ser pro lys ser lys pro phe ala
57          50          55          60
59 act gat agt ggc gcg atg tca cat aaa ctt gag aaa gct gac tta cta      240
60 thr asp ser gly ala met ser his lys leu glu lys ala asp leu leu
61 65          70          75          80

```

## RAW SEQUENCE LISTING

DATE: 04/23/2002

PATENT APPLICATION: US/09/940,235

TIME: 15:34:20

Input Set : A:\07064-009002.TXT

Output Set: N:\CRF3\04232002\I940235.raw

63	aag gct att caa gaa caa ttg atc gct aac gtc cac agt aac gac gac	288
64	Lys Ala Ile Gln Glu Gln Leu Ile Ala Asn Val His Ser Asn Asp Asp	
65	85 90 95	
67	tac ttt gag gtc att gat ttt gca agc gat gca acc att act gat cga	336
68	Tyr Phe Glu Val Ile Asp Phe Ala Ser Asp Ala Thr Ile Thr Asp Arg	
69	100 105 110	
71	aac ggc aag gtc tac ttt gct gac aaa gat ggt tcg gta acc ttg ccg	384
72	Asn Gly Lys Val Tyr Phe Ala Asp Lys Asp Gly Ser Val Thr Leu Pro	
73	115 120 125	
75	acc caa cct gtc caa gaa ttt ttg cta agc gga cat gtg cgc gtt aga	432
76	Thr Gln Pro Val Gln Glu Phe Leu Leu Ser Gly His Val Arg Val Arg	
77	130 135 140	
79	cca tat aaa gaa aaa cca ata caa aac caa gcg aaa tct gtt gat gtg	480
80	Pro Tyr Lys Glu Lys Pro Ile Gln Asn Gln Ala Lys Ser Val Asp Val	
81	145 150 155 160	
83	gaa tat act gta cag ttt act ccc tta aac cct gat gac gat ttc aga	528
84	Glu Tyr Thr Val Gln Phe Thr Pro Leu Asn Pro Asp Asp Asp Phe Arg	
85	165 170 175	
87	cca ggt ctc aaa gat act aag cta ttg aaa aca cta gct atc ggt gac	576
88	Pro Gly Leu Lys Asp Thr Lys Leu Lys Thr Leu Ala Ile Gly Asp	
89	180 185 190	
91	acc atc aca tct caa gaa tta cta gct caa gca caa agc att tta aac	624
92	Thr Ile Thr Ser Gln Glu Leu Leu Ala Gln Ala Gln Ser Ile Leu Asn	
93	195 200 205	
95	aaa aac cac cca ggc tat acg att tat gaa cgt gac tcc tca atc gtc	672
96	Lys Asn His Pro Gly Tyr Thr Ile Tyr Glu Arg Asp Ser Ser Ile Val	
97	210 215 220	
99	act cat gac aat gac att ttc cgt acg att tta cca atg gat caa gag	720
100	Thr His Asp Asn Asp Ile Phe Arg Thr Ile Leu Pro Met Asp Gln Glu	
101	225 230 235 240	
103	ttt act tac cgt gtt aaa aat cgg gaa caa gct tat agg atc aat aaa	768
104	Phe Thr Tyr Arg Val Lys Asn Arg Glu Gln Ala Tyr Arg Ile Asn Lys	
105	245 250 255	
107	aaa tct ggt ctg aat gaa gaa ata aac aac act gac ctg atc tct gag	816
108	Lys Ser Gly Leu Asn Glu Glu Ile Asn Asn Thr Asp Leu Ile Ser Glu	
109	260 265 270	
111	aaa tat tac gtc ctt aaa aaa ggg gaa aag ccg tat gat ccc ttt gat	864
112	Lys Tyr Tyr Val Leu Lys Lys Gly Glu Lys Pro Tyr Asp Pro Phe Asp	
113	275 280 285	
115	cgc agt cac ttg aaa ctg ttc acc atc aaa tac gtt gat gtc gat acc	912
116	Arg Ser His Leu Lys Leu Phe Thr Ile Lys Tyr Val Asp Val Asp Thr	
117	290 295 300	
119	aac gaa ttg cta aaa agt gag cag ctc tta aca gct agc gaa cgt aac	960
120	Asn Glu Leu Leu Lys Ser Glu Gln Leu Leu Thr Ala Ser Glu Arg Asn	
121	305 310 315 320	
123	tta gac ttc aga gat tta tac gat cct cgt gat aag gct aaa cta ctc	1008
124	Leu Asp Phe Arg Asp Leu Tyr Asp Pro Arg Asp Lys Ala Lys Leu Leu	
125	325 330 335	
127	tac aac aat ctc gat gct ttt ggt att atg gac tat acc tta act gga	1056

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/940,235

DATE: 04/23/2002

TIME: 15:34:20

Input Set : A:\07064-009002.TXT

Output Set: N:\CRF3\04232002\I940235.raw

```

128 Tyr Asn Asn Leu Asp Ala Phe Gly Ile Met Asp Tyr Thr Leu Thr Gly
129          340          345          350
131 aaa gta gag gat aat cac gat gac acc aac cgt atc ata acc gtt tat      1104
132 Lys Val Glu Asp Asn His Asp Asp Thr Asn Arg Ile Ile Thr Val Tyr
133          355          360          365
135 atg ggc aag cga ccc gaa gga gag aat gct agc tat cat tta gcc tat      1152
136 Met Gly Lys Arg Pro Glu Gly Glu Asn Ala Ser Tyr His Leu Ala Tyr
137          370          375          380
139 gat aaa gat cgt tat acc gaa gaa gaa cga gaa gtt tac agc tac ctg      1200
140 Asp Lys Asp Arg Tyr Thr Glu Glu Glu Arg Glu Val Tyr Ser Tyr Leu
141 385          390          395          400
143 cgt tat aca ggg aca cct ata cct gat aac cct aac gac aaa      1242
144 Arg Tyr Thr Gly Thr Pro Ile Pro Asp Asn Pro Asn Asp Lys
145          405          410
147 taa      1245
149 <210> SEQ ID NO: 2
150 <211> LENGTH: 414
151 <212> TYPE: PRT
152 <213> ORGANISM: Streptococcus equisimilis
154 <400> SEQUENCE: 2
155 Ile Ala Gly Pro Glu Trp Leu Leu Asp Arg Pro Ser Val Asn Asn Ser
156 1          5          10          15
157 Gln Leu Val Val Ser Val Ala Gly Thr Val Glu Gly Thr Asn Gln Asp
158          20          25          30
159 Ile Ser Leu Lys Phe Phe Glu Ile Asp Leu Thr Ser Arg Pro Ala His
160          35          40          45
161 Gly Gly Lys Thr Glu Gln Gly Leu Ser Pro Lys Ser Lys Pro Phe Ala
162          50          55          60
163 Thr Asp Ser Gly Ala Met Ser His Lys Leu Glu Lys Ala Asp Leu Leu
164 65          70          75          80
165 Lys Ala Ile Gln Glu Gln Leu Ile Ala Asn Val His Ser Asn Asp Asp
166          85          90          95
167 Tyr Phe Glu Val Ile Asp Phe Ala Ser Asp Ala Thr Ile Thr Asp Arg
168          100          105          110
169 Asn Gly Lys Val Tyr Phe Ala Asp Lys Asp Gly Ser Val Thr Leu Pro
170          115          120          125
171 Thr Gln Pro Val Gln Glu Phe Leu Leu Ser Gly His Val Arg Val Arg
172          130          135          140
173 Pro Tyr Lys Glu Lys Pro Ile Gln Asn Gln Ala Lys Ser Val Asp Val
174 145          150          155          160
175 Glu Tyr Thr Val Gln Phe Thr Pro Leu Asn Pro Asp Asp Asp Phe Arg
176          165          170          175
177 Pro Gly Leu Lys Asp Thr Lys Leu Leu Lys Thr Leu Ala Ile Gly Asp
178          180          185          190
179 Thr Ile Thr Ser Gln Glu Leu Leu Ala Gln Ala Gln Ser Ile Leu Asn
180          195          200          205
181 Lys Asn His Pro Gly Tyr Thr Ile Tyr Glu Arg Asp Ser Ser Ile Val
182          210          215          220
183 Thr His Asp Asn Asp Ile Phe Arg Thr Ile Leu Pro Met Asp Gln Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/940,235

DATE: 04/23/2002

TIME: 15:34:20

Input Set : A:\07064-009002.TXT

Output Set: N:\CRF3\04232002\I940235.raw

```

184 225          230          235          240
185 Phe Thr Tyr Arg Val Lys Asn Arg Glu Gln Ala Tyr Arg Ile Asn Lys
186          245          250          255
187 Lys Ser Gly Leu Asn Glu Glu Ile Asn Asn Thr Asp Leu Ile Ser Glu
188          260          265          270
189 Lys Tyr Tyr Val Leu Lys Lys Gly Glu Lys Pro Tyr Asp Pro Phe Asp
190          275          280          285
191 Arg Ser His Leu Lys Leu Phe Thr Ile Lys Tyr Val Asp Val Asp Thr
192          290          295          300
193 Asn Glu Leu Leu Lys Ser Glu Gln Leu Leu Thr Ala Ser Glu Arg Asn
194 305          310          315          320
195 Leu Asp Phe Arg Asp Leu Tyr Asp Pro Arg Asp Lys Ala Lys Leu Leu
196          325          330          335
197 Tyr Asn Asn Leu Asp Ala Phe Gly Ile Met Asp Tyr Thr Leu Thr Gly
198          340          345          350
199 Lys Val Glu Asp Asn His Asp Asp Thr Asn Arg Ile Ile Thr Val Tyr
200          355          360          365
201 Met Gly Lys Arg Pro Glu Gly Glu Asn Ala Ser Tyr His Leu Ala Tyr
202          370          375          380
203 Asp Lys Asp Arg Tyr Thr Glu Glu Glu Arg Glu Val Tyr Ser Tyr Leu
204 385          390          395          400
205 Arg Tyr Thr Gly Thr Pro Ile Pro Asp Asn Pro Asn Asp Lys
206          405          410
208 <210> SEQ ID NO: 3
209 <211> LENGTH: 777
210 <212> TYPE: DNA
211 <213> ORGANISM: Homo sapiens
213 <220> FEATURE:
214 <221> NAME/KEY: CDS
215 <222> LOCATION: (1)...(777)
217 <400> SEQUENCE: 3
218 cag gct cag caa atg gtt cag ccc cag tcc ccg gtg gct gtc agt caa      48
219 Gln Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val Ala Val Ser Gln
220 1          5          10          15
222 agc aag ccc ggt tgt tat gac aat gga aaa cac tat cag ata aat caa      96
223 Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr Gln Ile Asn Gln
224          20          25          30
226 cag tgg gag cgg acc tac cta ggt aat gtg ttg gtt tgt act tgt tat      144
227 Gln Trp Glu Arg Thr Tyr Leu Gly Asn Val Leu Val Cys Thr Cys Tyr
228          35          40          45
230 gga gga agc cga ggt ttt aac tgc gaa agt aaa cct gaa gct gaa gag      192
231 Gly Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro Glu Ala Glu Glu
232          50          55          60
234 act tgc ttt gac aag tac act ggg aac act tac cga gtg ggt gac act      240
235 Thr Cys Phe Asp Lys Tyr Thr Gly Asn Thr Tyr Arg Val Gly Asp Thr
236 65          70          75          80
238 tat gag cgt cct aaa gac tcc atg atc tgg gac tgt acc tgc atc ggg      288
239 Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys Thr Cys Ile Gly
240          85          90          95

```

## RAW SEQUENCE LISTING

DATE: 04/23/2002

PATENT APPLICATION: US/09/940,235

TIME: 15:34:20

Input Set : A:\07064-009002.TXT

Output Set: N:\CRF3\04232002\I940235.raw

```

242 gct ggg cga ggg aga ata agc tgt acc atc gca aac cgc tgc cat gaa      336
243 Ala Gly Arg Gly Arg Ile Ser Cys Thr Ile Ala Asn Arg Cys His Glu
244      100      105      110
246 ggg ggt cag tcc tac aag att ggt gac acc tgg agg aga cca cat gag      384
247 Gly Gly Gln Ser Tyr Lys Ile Gly Asp Thr Trp Arg Arg Pro His Glu
248      115      120      125
250 act ggt ggt tac atg tta gag tgt gtg tgt ctt ggt aat gga aaa gga      432
251 Thr Gly Gly Tyr Met Leu Glu Cys Val Cys Leu Gly Asn Gly Lys Gly
252      130      135      140
254 gaa tgg acc tgc aag ccc ata gct gag aag tgt ttt gat cat gct gct      480
255 Glu Trp Thr Cys Lys Pro Ile Ala Glu Lys Cys Phe Asp His Ala Ala
256 145      150      155      160
258 ggg act tcc tat gtg gtc gga gaa acg tgg gag aag ccc tac caa ggc      528
259 Gly Thr Ser Tyr Val Val Gly Glu Thr Trp Glu Lys Pro Tyr Gln Gly
260      165      170      175
262 tgg atg atg gta gat tgt act tgc ctg gga gaa ggc agc gga cgc atc      576
263 Trp Met Met Val Asp Cys Thr Cys Leu Gly Glu Gly Ser Gly Arg Ile
264      180      185      190
266 act tgc act tct aga aat aga tgc aac gat cag gac aca agg aca tcc      624
267 Thr Cys Thr Ser Arg Asn Arg Cys Asn Asp Gln Asp Thr Arg Thr Ser
268      195      200      205
270 tat aga att gga gac acc tgg agc aag aag gat aat cga gga aac ctg      672
271 Tyr Arg Ile Gly Asp Thr Trp Ser Lys Lys Asp Asn Arg Gly Asn Leu
272      210      215      220
274 ctc cag tgc atc tgc aca ggc aac ggc cga gga gag tgg aag tgt gag      720
275 Leu Gln Cys Ile Cys Thr Gly Asn Gly Arg Gly Glu Trp Lys Cys Glu
276 225      230      235      240
278 agg cac acc tct gtg cag acc aca tcg agc gga tct ggc ccc ttc acc      768
279 Arg His Thr Ser Val Gln Thr Thr Ser Ser Gly Ser Gly Pro Phe Thr
280      245      250      255
282 gat gtt cgt      777
283 Asp Val Arg
287 <210> SEQ ID NO: 4
288 <211> LENGTH: 259
289 <212> TYPE: PRT
290 <213> ORGANISM: Homo sapiens
292 <400> SEQUENCE: 4
293 Gln Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val Ala Val Ser Gln
294 1      5      10      15
295 Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr Gln Ile Asn Gln
296      20      25      30
297 Gln Trp Glu Arg Thr Tyr Leu Gly Asn Val Leu Val Cys Thr Cys Tyr
298      35      40      45
299 Gly Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro Glu Ala Glu Glu
300      50      55      60
301 Thr Cys Phe Asp Lys Tyr Thr Gly Asn Thr Tyr Arg Val Gly Asp Thr
302 65      70      75      80
303 Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys Thr Cys Ile Gly
304      85      90      95

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/940,235

DATE: 04/23/2002

TIME: 15:34:21

Input Set : A:\07064-009002.TXT

Output Set: N:\CRF3\04232002\I940235.raw

L:21 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:396 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:399 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:7